

Keiwa College English Placement Test Study 1995-1998

James B. Brown

Abstract

The progress of students in a skills-based, coordinated English language curriculum at Keiwa College is evaluated by the readministration of the placement test to students who have completed Level 3 in one or more of the skill areas. Students who start at Level 1 show the greatest improvement, while those who begin at Level 2 show improvement in some areas.

Introduction

In 1995 Keiwa College undertook a revision of the English language curriculum. The original curriculum of English language classes consisted of three "levels" of English divided into two course-types with different focuses respectively. One track focused on reading and writing while the other track emphasized conversation. The reading and writing classes met once a week for 90 minutes, while some of the conversation sessions met twice weekly for 45 minutes per session. Students entered the first classes in their first year and moved to the second and third levels in subsequent years upon successful completion of each course. There was no placement procedure implemented, so classes were of mixed levels. Teachers were free to choose their own textbooks and develop their own evaluation procedures for each of the sections of each course at the same "level."

The new core curriculum, in contrast, consists of three levels of courses in four skill areas: levels one through three in each of reading, writing, listening, and speaking. Students take a placement test upon entry into Keiwa College and are placed at the appropriate level in each skill area. Students may place in a higher level in some skills than others. All the teachers of

the various sections of each course use the same textbook and, with the exception of Speaking, skill-based tests are standardized across sections of each course. In other words, it is a coordinated curriculum.

The Placement Test used in the new curriculum is an in-house-developed placement instrument, consisting of a listening section with 30 problems (15 minutes), a reading section with four readings and 20 questions (45 minutes), a grammar section with 60 problems (30 minutes) and a writing sample based on a choice from among several topics (30 minutes). The whole test takes two hours.

In 1997 the first of the students who had entered the new curriculum in 1995 reached Level Three in various skill areas. The Language Curriculum Reform Committee decided to administer the relevant portions of the Placement Test to Level Three students and compare the results to those that they had achieved on the Placement Test that they had taken upon entry. This paper presents the results of that study.

The Data

Mean:	Std. Dev.:	Std. Error:	Variance:	Coef. Var.:	Count:
12.346	3.989	.782	15.915	32.313	26
Minimum:	Maximum:	Range:	Sum:	Sum Squared:	* Missing:
5	22	17	321	4361	2

Fig. 1: 1995 Listening Placement Test Scores of Those Who Reached Level 3 in 1997

Mean:	Std. Dev.:	Std. Error:	Variance:	Coef. Var.:	Count:
10.929	4.552	.304	20.721	41.653	224
Minimum:	Maximum:	Range:	Sum:	Sum Squared:	* Missing:
0	23	23	2448	31374	1

Fig. 2: 1995 Listening Placement Test Scores of All Students

The results shown in Fig. 1 are the scores on the 1995 Listening Placement Test of those students who were in Level 3 in 1997. There are 30 questions on the Listening Placement Test, so the mean was about 41%. The high score was 22 while the low score was five. If we compare the results of these students who eventually reached Level 3 to the class of '95 as a whole (Fig. 2), we can see that they were slightly more proficient with a mean score of a little more than one point (1.417) higher. The high score of the whole class was a point higher than those who went on to Level 3.

A statistical comparison of these two groups shown below (Fig. 3) indicates that there is no significant statistical difference between the whole group of entering freshmen in 1995 and those among them who went on to take Level 3 classes.

Unpaired t-Test					
DF:	X Count:	Y Count:	Mean X:	Mean Y:	Unpaired t Value:
248	26	224	12.346	10.929	1.521
.05 < p ≤ .1					

Fig. 3: Comparison of Listening Placement Test Scores of All 1995 Freshmen and Those Who Went to Level 3 in 1997

Since no more than a .05 level of significance (p in the data display) is required to demonstrate a statistically significant difference (95% probability that the results did not happen by chance), the levels shown in the table above indicate a lack of statistically significant difference between the two groups.

Having demonstrated that students who went on to Level 3 were not significantly different in listening skills in 1995 from their peers who did not, we can compare their initial scores on the Listening Placement Test to those they received when they took the test again while they were in Level 3.

Mean:	Std. Dev.:	Std. Error:	Variance:	Coef. Var.:	Count:
22.429	3.048	.576	9.291	13.59	28
Minimum:	Maximum	Range:	Sum:	Sum Squared:	* Missing:
15	27	12	628	14336	0

Fig. 4: Listening Placement Results of 1997 Level 3 Listening Students Who Entered at Level One in 1995

A quick look at the mean in Fig. 4 indicates that progressing from Level 1 to Level 3 in listening resulted in an improvement of about 10 points (33%) on the Placement Test.

Statistical comparison (Fig. 5) indicates the obvious: that there is about a 5 in 10,000 chance that these results could have happened at random — a highly significant finding.

Paired t-Test		
DF:	Mean X - Y:	Paired t value:
25	10.192	11.506
$p \leq .0005$		
Note: 2 case(s) deleted with missing values.		

Fig. 5: Comparison Listening Scores 95/97

Some of the students taking the Listening Placement Test in 1997 placed into Level 2 Listening in 1996. We thought it would be interesting to see if there was as dramatic an improvement in their scores after only two years of classes as there was for those who placed in Level One in 1995, a year earlier.

Mean:	Std. Dev.:	Std. Error:	Variance:	Coef. Var.:	Count:
18.526	4.477	1.027	20.041	24.164	19
Minimum:	Maximum:	Range:	Sum:	Sum Squared:	* Missing:
10	25	15	352	6882	0

Fig. 6: 1996 Listening Placement Test of Those Who Reached Level 3 in 1997

A comparison of Fig. 6 with Fig. 1 shows the difference in mean scores between those who place in Level 1 and those

who place in Level 2. The score of those who placed into Level 2 in 1996 is about 6 points, or about 20 percent, higher than those who placed into Level 1 in 1995.

Comparing this group's scores with those they achieved on the Placement Test in 1997 (Fig. 7) shows considerable improvement with a higher mean than those students who started in 1995 (Fig. 4). The gap between the first test scores and those taken in 1997 is not as high, however, only 6.1 points or about 20 percent.

Mean:	Std. Dev.:	Std. Error:	Variance:	Coef. Var.:	Count:
24.6	2.633	.833	6.933	10.704	10
Minimum:	Maximum:	Range:	Sum:	Sum Squared:	* Missing:
20	29	9	246	6114	9

Fig. 7: Listening Placement Results of 1997 Level 3 Listening Students Who Entered at Level Two in 1996

Nevertheless, a statistical analysis of the students' scores (Fig. 8) shows a highly significant improvement.

Paired t-Test		
DF:	Mean X - Y:	Paired t value:
9	2.8	2.905
$.005 < p \leq .01$		
Note: 9 case(s) deleted with missing values.		

Fig. 8: Comparison Listening 96/97¹

¹ The mean difference shown in Fig. 8 is different from that shown in Figs. 7 and 6. This is due to the missing values.

What about the scores on the Reading portion of the Placement Test? The Reading portion of the Placement Test consists of 20 questions with four readings of increasing difficulty.

The first step is to determine whether there was any significant difference between students who did not go on to Level 3 and those who did at the time of the 1995 Placement Test taken when they all entered Keiwa College. Figs. 9 through

11 show the results of that comparison. Fig. 11 indicates that there was no statistically significant difference between the two groups on the Reading Placement Test.

Mean:	Std. Dev.:	Std. Error:	Variance:	Coef. Var.:	Count:
7.527	2.741	.37	7.513	36.414	55
Minimum:	Maximum:	Range:	Sum:	Sum Squared:	* Missing:
0	14	14	414	3522	0

Fig. 9: 1995 Reading Placement Test Scores of Those Who Reached Level 3 in 1997

Mean:	Std. Dev.:	Std. Error:	Variance:	Coef. Var.:	Count:
7.371	3.188	.213	10.163	43.252	224
Minimum:	Maximum:	Range:	Sum:	Sum Squared:	* Missing:
0	14	14	1651	14435	1

Fig. 10: 1995 Reading Placement Test Scores of All Students

Unpaired t-Test					
DF:	X Count:	Y Count:	Mean X:	Mean Y:	Unpaired t Value:
277	224	55	7.371	7.527	-.335
.1 < p < .375					

Fig. 11: Comparison of Reading Placement Test Scores of All 1995 Freshmen and Those Who Went to Level 3 in 1997

As with the Listening Placement Test, the Reading Placement Test was administered to students in Level 3 Reading classes to determine how much, if any, progress had been made during the course of their study from Level 1.

A quick glance at Fig. 12 reveals that the mean for the test given in 1997 has jumped more than 3 points or about 15 percent. Fig. 13 indicates that this increase is highly significant.

Mean:	Std. Dev.:	Std. Error:	Variance:	Coef. Var.:	Count:
10.667	2.747	.44	7.544	25.749	39
Minimum:	Maximum:	Range:	Sum:	Sum Squared:	# Missing:
6	16	10	416	4724	16

Fig. 12: Reading Placement Results of 1997 Level 3 Reading Students Who Entered at Level One in 1995

Paired t-Test		
DF:	Mean X - Y:	Paired t value:
38	3.231	6.196
$p \leq .0005$		
Note: 16 case(s) deleted with missing values.		

Fig. 13: Comparison Reading Scores 95/97*

* The discrepancy in mean differentials is due to missing values.

As with the Listening Placement Test, we studied the Reading scores for those students who entered Level 2 Reading in 1996. Figs. 14 through 16 show the results of this investigation.

Mean:	Std. Dev.:	Std. Error:	Variance:	Coef. Var.:	Count:
11.895	2.514	.577	6.322	21.138	19
Minimum:	Maximum:	Range:	Sum:	Sum Squared:	# Missing:
6	17	11	226	2802	0

Fig. 14: 1996 Reading Placement Test of Those Who Reached Level 3 in 1997

Mean:	Std. Dev.:	Std. Error:	Variance:	Coef. Var.:	Count:
11.846	2.478	.687	6.141	20.919	13
Minimum:	Maximum:	Range:	Sum:	Sum Squared:	# Missing:
9	17	8	154	1898	6

Fig. 15: Reading Placement Results of 1997 Level 3 Reading Students Who Entered at Level Two in 1996

Paired t-Test		
DF:	Mean X - Y:	Paired t value:
12	-1.154	-1.376
.05 < p ≤ .1		
Note: 6 case(s) deleted with missing values.		

Fig. 16: Comparison Reading 96/97#

* The discrepancy in mean differentials is due to missing values.

Fig. 16 indicates that there is no significant difference between the scores achieved by the students in 1996 and those achieved by the same students in 1997. Careful readers will have noticed that the mean of the Level 2 placement scores for 1996 are actually higher than those of students who entered in 1995 and took the test at Level 3 in 1997. Statistical comparison of these two scores, however, indicates no significant difference between the two groups. (Please refer to the Discussion part of this paper for further information about the meaning of these scores.)

Finally, what about the scores on the Grammar Placement Test? The Grammar Placement Test consists of 60 items. In 1997 it was administered to students in Level 3 Writing classes. First the Grammar scores of those who went on to Level 3 Writing were compared to the scores of those who did not. Since there is almost a 3-point gap between the scores of the two groups the statistical comparison is shown below in Fig. 19. It demonstrates that although the group that went on to Level 3 Writing scored slightly higher, there is no statistically significant difference between the score groups. Differences could be accounted for by random variations.

Mean:	Std. Dev.:	Std. Error:	Variance:	Coef. Var.:	Count:
37.564	11.805	1.592	139.362	31.427	55
Minimum:	Maximum:	Range:	Sum:	Sum Squared:	* Missing:
0	54	54	2066	85132	0

Fig. 17: 1995 Grammar Placement Test Scores of Those Who Reached Level 3 in 1997

Mean:	Std. Dev.:	Std. Error:	Variance:	Coef. Var.:	Count:
34.759	11.611	.776	134.812	33.404	224
Minimum:	Maximum:	Range:	Sum:	Sum Squared:	* Missing:
0	56	56	7786	300696	1

Fig. 18: 1995 Grammar Placement Test Scores of All Students

Unpaired t-Test					
DF:	X Count:	Y Count:	Mean X:	Mean Y:	Unpaired t Value:
277	224	55	34.759	37.564	-1.6
.05 < p ≤ .1					

Fig. 19: Comparison of Grammar Placement Test Scores of All 1995 Freshmen and Those Who Went to Level 3 in 1997

Next the test scores achieved on the Placement Test given in 1997 (Fig. 20) were compared to those achieved in 1995. The mean score increased from 37.564 to 46, an increase of more than 8 points or 13 percent. Fig. 21 shows that this increase cannot be accounted for by random factors and is statistically highly significant.

Mean:	Std. Dev.:	Std. Error:	Variance:	Coef. Var.:	Count:
46	7.15	1.209	51.118	15.543	35
Minimum:	Maximum:	Range:	Sum:	Sum Squared:	* Missing:
26	57	31	1610	75798	20

Fig. 20: Grammar Placement Results of 1997 Level 3 Writing Students Who Entered at Level One in 1995

Paired t-Test		
DF:	Mean X - Y:	Paired t value:
34	8.457	4.709
p ≤ .0005		
Note: 20 case(s) deleted with missing values.		

Fig. 21: Comparison Grammar Scores 95/97⁺

⁺ The discrepancy in mean differentials is due to missing values.

As with the Listening and the Reading Placement Test, the Grammar Placement Test results of students who placed into Level 2 Writing in 1996 (Fig. 22) were compared to their scores on the test in 1997 (Fig 23) when they were in Level 3 Writing.

Mean:	Std. Dev.:	Std. Error:	Variance:	Coef. Var.:	Count:
50.75	1.893	.946	3.583	3.73	4
Minimum:	Maximum:	Range:	Sum:	Sum Squared:	* Missing:
48	52	4	203	10313	0

Fig. 22: 1996 Grammar Placement Test of Those Who Reached Level 3 in 1997

Mean:	Std. Dev.:	Std. Error:	Variance:	Coef. Var.:	Count:
52.75	2.5	1.25	6.25	4.739	4
Minimum:	Maximum:	Range:	Sum:	Sum Squared:	* Missing:
50	56	6	211	11149	0

Fig. 23: Grammar Placement Results of 1997 Level 3 Writing Students Who Entered at Level Two in 1996

Paired t-Test		
DF:	Mean X - Y:	Paired t value:
3	2	.926
.1 < p ≤ .375		

Fig. 24: Comparison Grammar 96/97

The results of the comparison (Fig. 24) show that there was no statistically significant difference between the scores achieved in 1996 and those reached in 1997.

Discussion

The results of the study suggest that students entering at Level 1 in Listening show marked improvement by the time they reach Level 3. There are several factors that might play a role in this improvement. One factor could be that initial levels

were so low that any improvement would skew statistical evaluation. Indeed, with an initial mean score of 12.346 points out of 30 on the 1995 placement test, a score of about 41 percent, this might seem to be a contributing factor. Nevertheless, a ten-point increase in the score on the test suggests that more than simple mathematical artifacts are at work here. Especially when one examines the data from the group who placed into Level 2 in 1996 and how they fared in 1997, improvement due to teaching and the new curriculum seem to be the significant factor. Starting at about 62%, they increased their mean score to 82% in less than two years. Improvements in listening comprehension may also be influenced by the fact that not only all the Listening and Speaking classes, but also many of the Reading and Writing classes are taught in English. This, combined with the fact that many other classes outside of the core-curriculum are also taught in English, may be the most important factor in skill improvement.

Interestingly, in a questionnaire of 4th-year students conducted in 1998 by Profs. Kanayama and Williams (Kanayama/Williams 1998) on student perceptions of language-skill progress, improvement in listening at Level 1 was cited as the highest (among all skill areas) with a gradual dropping off in Level 2 and a further decline in Level 3. Students themselves thought that progress was greatest in the earlier levels of listening.

Reading scores by comparison do not show similar levels of improvement despite the fact that initial scores in 1995 were slightly lower than those achieved on the Listening Placement Test — 37.6% compared to 41.2%.

There are several factors which might have played a role in the production of these figures. The most important influencing element is probably the fact that one of the problems on the Reading Placement Test was changed in 1996. While the overall number of questions did not change, the fact that one of the readings was replaced might suggest that it was a substantially different test. A comparison of the overall scores between the

1995 Placement Test and the 1996 Placement Test, however, shows no statistical difference between the two tests. However, no comparison of how students would have scored on the modified Reading Placement Test in 1995 is possible, so we cannot know for sure if the change in the test might be responsible for how the scores turned out.

The fact that students taking the test in 1996 and again in 1997 showed a small drop in their scores suggests that the test had become somewhat more difficult. This in turn would account for the fact that students placing in Level 1 in 1995 would not show the dramatic improvements in Reading that they did in Listening.

The Grammar test on the other hand shows the same trends as the Listening test. Initial mean scores of 62.6% in 1995 were brought up to 76.7% by 1997. Students entering Level 2 Writing in 1996 scored a mean of 84.6% on the Grammar Placement Test and increased that score to 87.9%. As the data shows, however, this increase is not statistically significant. As with very low scores showing enormous improvement due to small mathematical increases, rather high scores remain statistically unaffected by similar increments. In addition, the relatively low numbers of those who took the test in 1997 may be an influencing factor.

One interesting feature of the data that should not be overlooked is the differences in Standard Deviation (Std. Dev. on the data displays) from the scores achieved by incoming freshmen in 1995 and those they achieved later in 1997. Even more conspicuous is the Standard Deviation of those who placed into Level 2 Writing in 1996, and how they fared in 1997. Standard Deviation indicates how far away from the mean scores are clustered. The higher the Standard Deviation, the greater the range of proficiency among the test takers. A narrowing of the Standard Deviation indicates that the students were all coming to be able to perform at similar levels of proficiency in each group.

Conclusion

Careful statistical analysis of student performance is a necessary part of any language program. Regrettably no such evaluation was carried out during the period of the more traditional English language curriculum at Keiwa College that preceded curriculum reform in 1995. Therefore we do not have a body of data that would allow us to compare the relative advantages or disadvantages of implementing a coordinated curriculum. In language programs where students are not initially placed by testing, this kind of year-by-year assessment is very difficult. Aside from placement testing, very little standardized examination is attempted on student bodies as a whole in many college-level language programs.

Which brings us to the reliability of the Keiwa Placement Test. In retrospect, it would have been more reliable to have instituted a well-established testing instrument such as an Institutional TOEFL test or the Michigan Test instead of developing an in-house placement procedure. Modeled on other tests, including placement tests used in Intensive English Programs in the US, the Keiwa Placement Test did not undergo the rigorous testing necessary to develop high levels of confidence in its results. The alteration of the Reading portion of the test in 1996 seemed to have had some effect on the scores students were able to achieve. Nevertheless, the bulk of it remained unchanged and was tested without implementation in 1994 and has been used since then. Cross comparisons of results from 1995 to 1998, the most recent results, indicate that there are no statistically significant differences in placement scores for any of the three areas shown in this paper. Internal consistency, at least, does not seem to be a major problem.

In the future, it might be good to move away from the in-house placement procedure to a tested language proficiency instrument. In this way, analysis of year-to-year performance and improvements can be carried out with the assurance that data retrieved is accurate enough to be used with confidence.

References

Kanayama, A. & Williams. J. *Questionnaire of 4th-year Students on the Language Program*. (Keiwa College) July 1998.